JOURNAL OF BACTERIOLOGY, Mar. 1993, p. 1264–1271
0021-9193/93/051264-08$02.00/0

Vol. 175, No. 5

Implications of Tn*5*-Associated Adjacent Deletions

ROSS A. JILK, JOHN C. MAKRIS,† LYNNE BORCHARDT,‡ AND WILLIAM S. REZNIKOFF*

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin—Madison, Madison, Wisconsin 53706

Received 1 September 1992/Accepted 23 December 1992

The prokaryotic transposable element Tn*5* has been found to promote the formation of adjacent deletions. The frequency of adjacent deletion formation is much lower than that of normal transposition events. Like normal transposition, however, adjacent deletion formation requires the activity of the transposase protein. The deletions can be divided into two classes, as distinguished by their endpoints. The occurrence of one of the two deletion classes is increased when the frequency of normal transposition is reduced by the introduction of a deletion or a certain base substitution at one of the two outside ends (OEs). The nature of the base substitution at the mutant OE influences the class of deletion found adjacent to the wild-type OE, even though these two ends are about 12 kbp apart. By studying the formation of these deletions, we have gained some insight into the way in which the transposase interacts with the OEs. Our observations suggest that there is a protein-mediated interaction between the two ends, that different end base pairs are involved in different transposition-related processes, and that the adjacent deletions are the result of nonproductive attempts at transposition.

Transposons are a diverse class of mobile genetic elements that can promote genetic rearrangements without a requirement for sequence homology (5). While insertions are the predominant event, deletions and inversions of neighboring host DNA can also be found (2, 26). Adjacent deletions are a specific class of rearrangements in which the deletion of host DNA begins immediately next to one end of a transposable element. We have found that for Tn*5*, these adjacent deletions appear to result from failed attempts at transposition. An analysis of the reasons for the failure of these events may improve our understanding of the underlying mechanism of transposition.

Tn*5* is a composite bacterial transposon that consists of two IS*50* insertion sequences in an inverted orientation and flanking the gene for neomycin resistance (for a review of Tn*5*, see reference 4). IS*50*R encodes the functional *cis*-acting transposase protein (Tnp) and its inhibitor (Inh). The IS*50* and Tn*5* mobile elements are defined by the unique DNA sequences at their respective ends. IS*50* is flanked by two *cis*-acting 19-bp DNA sequences that share some sequence similarity and are called the outside end (OE) and the inside end (IE) (Fig. 1). Tn*5*, on the other hand, is bordered by two OE sequences. The integrity of the end sequences is necessary for normal levels of transposition (18, 24). In fact, recent data suggest that these ends contain the Tnp binding site (8). Each end is also involved in host-protein interactions; the OE contains a DnaA consensus binding site (32), and the IE contains two Dam methylation sites (19, 31) and overlaps a Fis binding site (30) (Fig. 1).

Tn*5* is speculated to transpose by a conservative mechanism (4). This mechanism of transposition involves excising the transposon from the donor site and inserting it into a new DNA target site without replication. Until recently, however, the detection of adjacent deletions has been taken as presumptive evidence of a replicative mechanism of transposition, because adjacent deletions have been thought to arise only through a cointegrate intermediate (26). Both Isberg and Syvanen (12) and Tomcsanyi et al. (28) have shown evidence for the ability of Tn*5* to promote adjacent deletions by insertion of Tn*5* into a site within the transposon itself. An alternative model that has recently been proposed to explain adjacent deletion formation involves the possibility that two transposable elements on sister chromosomes can each provide a single end to be used in transposition (17, 25). It has also been suggested that adjacent deletions might arise through aborted transposition reactions (17, 25, 28).

We have found two classes of adjacent deletions that are transposase dependent but can occur without a normal insertional transposition. At least one of these deletion classes is likely to be the product of an aborted transposition event. Various point mutations in one of the 19-bp OEs of Tn*5* reduce transposition to very low levels (18, 24). Throughout this work we use the word transposition to refer specifically to the insertion of the transposon into a new target site. These OE base substitutions allowed us to isolate deletions that occurred adjacent to the wild-type (WT) OE, located 12 kbp away. These OE base substitutions are related to the class of deletions isolated. The relationship between mutation type and deletion class permits us to propose a model that addresses the way in which the transposase binds to the OE and how the two OEs interact in a transposase-dependent fashion.

MATERIALS AND METHODS

Bacterial strains and plasmids. All clonings, papilla formation studies, and deletion isolations were done with *Escherichia coli* DH5α [*supE44* Δ*lacU169* (ϕ80*lacZ* ΔM15) *hsdR17 recA1 end A1 gyrA96 thi-1 rel A1*].

The construction of pRZ1490, pRZ1495, and all of the pRZ1495 derivatives carrying OE mutants has been described by Makris and coworkers (18, 23). Plasmid derivatives that contain an OE substitution are denoted with the following nomenclature: plasmid:base substitution (for example, pRZ1495:6C is pRZ1495 carrying a T→C change at

* Corresponding author.
† Deceased.
‡ Present address: Department of Veterinary Science, University of Wisconsin—Madison, Madison, WI 53706.

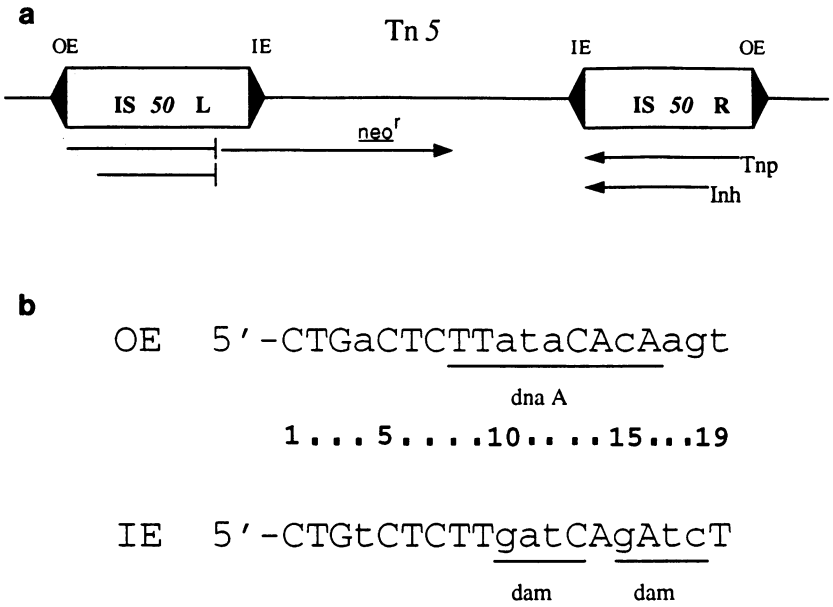

FIG. 1. (a) Composite transposon Tn*5*. Tn*5* is composed of two IS*50* sequences in an inverted orientation around a neomycin resistance gene. IS50_R encodes the proteins active in transposition and its regulation (Tnp and Inh). IS50_L contains a single base substitution that causes a premature stop codon that produces inactive forms of Tnp and Inh. This same base substitution, however, forms part of a promoter for the resistance gene. (b) Base sequences of the OE and the IE. Uppercase letters are used to highlight the positions that have identical bases at each end. IS*50* transposition requires an OE and an IE in inverted orientations relative to each other. Tn*5* transposition requires two inverted OEs. The OE contains a DnaA consensus binding site, and the IE contains two sites subject to Dam methylation. Numbers are in base pairs.

position 6 of the cassette OE [cOE] [see below]). Tnp$^-$ Inh$^-$ plasmids pRZ8520, pRZ8520:2A, and pRZ8520:6C were constructed by digestion of pRZ1495 and its pertinent end mutant derivatives with *Eag*I. The digests were then religated to form a 477-bp deletion. Tnp$^-$ Inh$^+$ plasmids pRZ8530:2A and pRZ8530:6C were constructed by isolation from pRZ407 of the *Eag*I fragment that contains a 17-bp deletion causing a frameshift mutation near the 5′ end of the Tnp reading frame (14) and cloning of that fragment into pRZ8520:2A and pRZ8520:6C digested with *Eag*I. Plasmid pRZ8595 contains a unique *Apa*I recognition site between the deletion-associated OE (dOE) (see below) and the 3′ end of the ColE1 *imm* gene. pRZ8595 was constructed by partial *Eco*RI digestion of pRZ1495 and insertion of a synthetic 16-bp oligonucleotide containing the *Apa*I recognition site (see Fig. 2).

Isolation of deletions. Derivatives of the pRZ1495 papilla formation vector were used to isolate deletions. pRZ1495 carries the Tn*5* derivative Tn*5*ORF*lac*, which contains the *lacZYA* genes but with the transcriptional and translational start sites of *lacZ* replaced by a complete OE sequence (Fig. 2) such that β-galactosidase is not expressed (15). This OE will be referred to as the dOE for clarity. DNA rearrangements that fuse the dOE in frame to an actively expressed gene can be detected by virtue of the resulting Lac$^+$ phenotype, as described by Krebs and Reznikoff (15). A *Kpn*I-*Sph*I cassette was used to introduce various base substitutions at the other OE (18). This OE will be referred to as the cOE (Fig. 2). All OE mutations discussed are at the cOE, while the dOE is always WT.

pRZ1495 derivatives containing a mutation in the cOE were introduced into DH5α and plated on lactose-MacConkey agar to yield about 50 colonies per plate. After 5 days at 37°C, the frequency of papilla formation was determined as the ratio of the number of papillae observed to the total number of colonies plated (pap/col). We isolated plasmids from papilla-producing cells and transformed them into competent DH5α cells. The transformants were then plated on indicator plates to select for plasmids containing Lac$^+$ fusions. The plasmids from Lac$^+$ transformants were reisolated. Only one Lac$^+$ isolate was analyzed for each papilla. Restriction analysis of the plasmids was used to determine whether the fusion event was a deletion, a plasmid-to-plasmid transposition, or some other plasmid rearrangement. The endpoints of the deletions were determined by sequencing of the dOE fusion. The observed deletion frequency was expressed as the ratio of the number of Lac$^+$ fusions due to deletions to the total number of Lac$^+$ papillae analyzed (del/pap). A deletion frequency of 1 means that deletions were the only papilla-generating events found. A deletion frequency of <1 indicates that the cOE in question was able to support some other fusion events, such as insertional transposition into the chromosome. The intrinsic deletion frequency (deletions per colony [del/col]) was calculated by multiplying pap/col by del/pap.

In the presence of two WT OEs, deletions are a very small fraction of the Lac$^+$ fusion-generating rearrangements. These deletions were found by use of pRZ1495 derivative pRZ8595. Pools of cells containing different Lac$^+$ fusions derived from pRZ8595 were cultured, and their plasmids were isolated. The plasmids were subjected to *Apa*I digestion before transformation to select for plasmids in which this site was deleted. Again, Lac$^+$ transformants were analyzed.

Isolation of Tn*5* transpositions into the *imm* gene. We isolated insertions of our Tn*5* derivative into the *imm* gene by using a protocol that screened cells from papillae for the loss of immunity to colicin. This procedure was done by plating a lawn of colicin-producing cells on a tryptone-yeast extract plate, killing the lawn of cells with chloroform

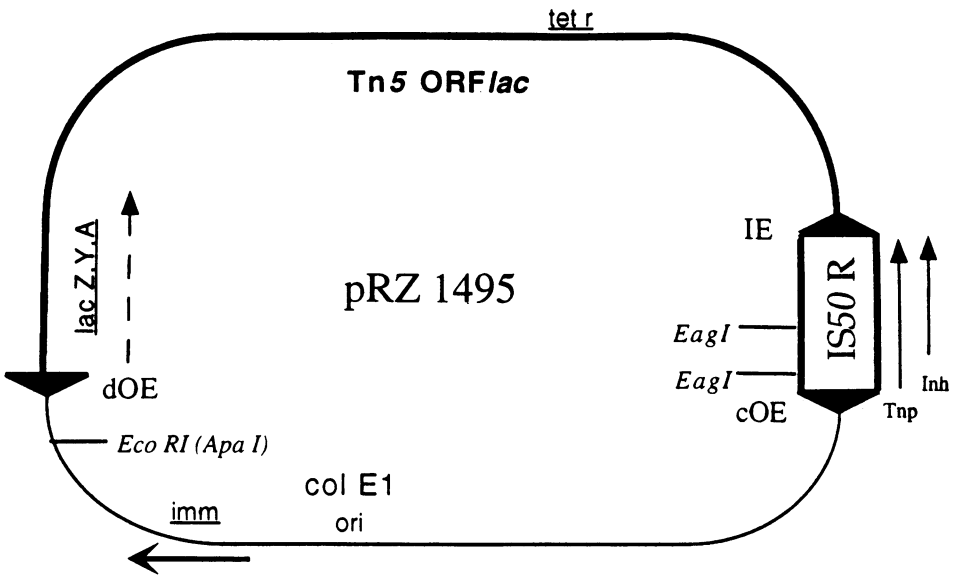

FIG. 2. Papillation vector pRZ1495. The structural sequences of $IS50_L$ have been replaced by *lacZYA* genes. The *lac* promoter and *lacZ* translational start sites have been deleted. The $IS50_L$ OE is still intact but is now called the dOE. A fusion of the dOE in frame to an active gene is required for the expression of β-galactosidase. A tetracycline resistance gene has been inserted as the antibiotic determinant. The $IS50_R$ OE has been replaced by an oligonucleotide cassette that contains the OE sequences and is now called the cOE. This procedure was done to allow the introduction of base substitutions into the end. The ColE1 *imm* gene is shown to emphasize its proximity and orientation relative to the dOE and *lac* sequences. The *Eco*RI site used to insert the *Apa*I oligonucleotide for the construction of pRZ8595 is shown as well.

vapors, overlaying the lysate with top agar, and then spotting Lac^+ cells on the top agar (29). Ten Lac^+ Imm^- insertions were sequenced to determine their insertion points.

Media, enzymes, and chemicals. *E. coli* liquid cultures were grown at 37°C in Luria broth. Luria broth medium, lactose-MacConkey agar, and minimal agar M9 were made as described by Miller (22). Medium preparations were purchased from Difco. Lactose was added to minimal media to 0.2%. Media used for selections contained tetracycline at 15 μg/ml and/or kanamycin at 40 μg/ml. Tetracycline and kanamycin were purchased from Sigma. The restriction enzymes *Apa*I, *Eag*I, and *Eco*RI and T4 DNA ligase were purchased from New England Biolabs. The sequencing enzyme Sequenase was purchased from U.S. Biochemical Corp.

RESULTS

We found two classes of transposase-dependent adjacent deletions. Class I deletions started immediately adjacent to the dOE and extended to a single target site. These deletions were found in conjunction with a subclass of mutations at the cOE (Fig. 3). The frequency of appearance of class I deletions appeared to be a property of the cOE mutations with which they were isolated, ranging from 0.2 del/col after 5 days for a 10A→T substitution to 6 del/col for a 6T→C substitution (Table 1). Class II deletions started 1 base outside of the dOE and extended to a variety of sequences. These deletions occurred in conjunction with a different subclass of cOE mutations (Fig. 3). The frequency of appearance of these deletions was ca. 0.03 del/col (Table 1). Both of these classes of deletions could also be isolated from an element containing WT OE sequences at both ends.

Class I deletions. The distinguishing feature of class I deletions is that they always extend to one target site. The deletions in this class begin immediately 5′ to the dOE and extend exactly 465 bp to the TAGAAT sequence in the ColE1 *imm* gene (Fig. 4) (7). This target site sequence shares

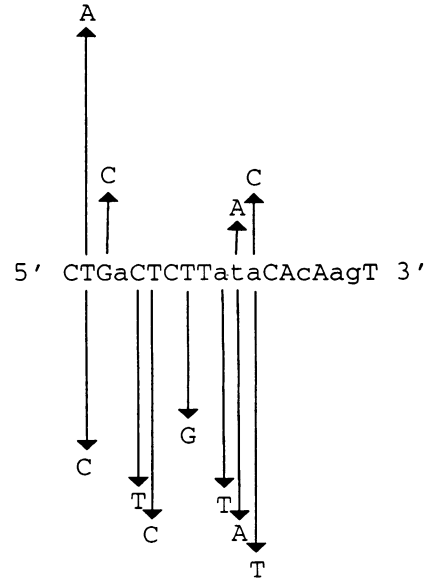

FIG. 3. Visual summary of cOE mutations studied and their associated deletion classes. Arrows pointing down represent mutations that led to class I deletions. Arrows pointing up represent mutations that led to class II deletions. The lengths of the arrows are scaled to represent the number of deletions isolated in association with a particular mutation.

TABLE 1. Characteristics of adjacent deletions

Class	End type	pRZ plasmid	No. of deletions studied	pap/col	del/pap	del/col
WT	WT	1495 and 8595	5 cII and 3 cI II	>100[a]	<0.0005[a]	0.05[a]
I	Deletion	1490	8	0.25	1.0	0.25
	2T→C	1495:2C	5	5	1.0	5
	5C→T	1495:5T	6	4	1.0	4
	6T deletion	1495:del6	4	0.25	1.0	0.25
	6T→C	1495:6C	8	6	1.0	6
	8T→G	1495:8G	4	3	1.0	3
	10A→T	1495:10T	6	0.20	1.0	0.20
	12A→T	1495:12T	8	5	1.0	5
II	2T→A	1495:2A	7	0.03	1.0	0.03
	3G→C	1495:3C	2	6	0.01	0.06
	12A→C	1495:12C	2	3	0.01	0.03
I and II	11T→A	1495:11A	7 cII and 1 cIII	4	1.0	4
III	WT	8520 (Tnp^- Inh^-)	6	0.01	1.0	0.01
	2T→A	8520:2A (Tnp^- Inh^-)	2	0.01	1.0	0.01
	6T→C	8520:6C (Tnp^- Inh^-)	2	0.01	1.0	0.01
	2T→A	8530:2A (Tnp^- Inh^+)	4	0.01	1.0	0.01
	6T→C	8530:6C (Tnp^- Inh^+)	2	0.01	1.0	0.01

[a] Plasmid pRZ1495 has a WT cOE, so the number of papillae per colony was too large to count after 5 days and the method of deletion isolation did not lend itself to determining deletion frequency. However, we were able to qualitatively note that we isolated a deletion from about 1 of every 20 overpapillated colonies studied, yielding a del/col of 0.05. Assuming that there are 100 papillae per colony, we can also estimate a del/pap of about 0.0005 with a WT cOE.

identity at five of six positions with bp 5 to 10 of the OE (Fig. 1).

Class I deletions occurred at a very low frequency when both OEs contained the WT sequence. Their frequency of occurrence was enhanced, however, by the introduction of certain mutations into the cOE. For example, when the cOE was functionally deleted by replacement with a *kan* resistance cassette (pRZ1490), class I deletions were the only papilla-generating events found, appearing at a frequency of 0.25 del/col (Table 1). Other cOE base substitutions that generated class I deletions, along with their frequencies, are shown in Table 1. All of the class I deletion-generating cOE mutations resulted in a deletion frequency of 1 del/pap. This means that none of these cOE mutations supported any observable level of transposition.

The role of the transposase protein in promoting class I deletions was studied by use of two types of mutations in the transposase gene. We made Tnp^- Inh^- constructs by removing the *Eag*I fragment from plasmids with WT and 6T→C cOEs. No class I deletions were found with these constructs. The only Lac^+ fusions that were found fell into a new deletion class with apparently random start and stop points. Their frequency of appearance was about 0.01 del/col. These nonspecific, Tnp-independent deletions will be referred to as class III deletions (Table 1). We next made a Tnp^- Inh^+ mutation to determine whether Inh alone could cause class I deletions. This mutation was made by creating a 17-bp frameshift deletion in the Tnp reading frame on a plasmid with a 6T→C cOE. All of the deletions that we isolated with this Tnp^- Inh^+ mutation were of class III. These results indicate that functional Tnp is required for the occurrence of class I deletions.

Class II deletions. The distinguishing feature of class II deletions is the unexpected position of their start point. All of the deletions in this class start precisely 1 base 5′ to the dOE (a single example is shown in Fig. 4). In addition, the endpoints of the class II deletions are at various positions in the ColE1 *imm* gene (Fig. 5). An inspection of the various class II deletion endpoints shown in Fig. 5 shows that for the large majority, we can confirm that the extra guanine originated from the −1 position. However, in three cases, the guanine could in theory have been derived from the deletion endpoint rather than from the −1 position.

Class II deletions were isolated at a low frequency in the presence of a WT cOE. These deletions were also found in conjunction with a particular set of point mutations at the cOE (Fig. 3). The frequency of class II deletion formation was low (0.03 to 0.06 del/col) for all the mutant and WT cOEs (Table 1). The method of isolating WT cOE deletions involved the use of colonies with indeterminate numbers of papillae, so the deletion frequency could not be reliably calculated. Qualitatively, however, we noted that we were able to isolate one WT cOE deletion from about 1 of every 20 overpapillated colonies, yielding a del/col of approximately 0.05.

Most of the class II-associated cOE mutations also supported some level of transposition, as shown by their observed deletion frequency of <1 del/pap. The exception was the 2T→A cOE mutation, which abolished all transposition and only gave rise to deletions. In fact, papilla formation frequencies associated with the 2T→A mutation were so low that several of the random, Tnp-independent class III deletions were also isolated.

The cOE mutation consisting of an 11T→A base substitution was the only cOE mutation that gave rise to class I (7) and class II deletions (1), as shown in Table 1.

The formation of class II deletions also required a functional transposase. The 2T→A cOE mutation was introduced into the Tnp^- Inh^- and Tnp^- Inh^+ constructs, and again only the random class III deletions were isolated.

Insertion events. We examined whether there was also a preferred target site for transposition in the *imm* gene. We isolated 10 Lac^+ insertions that were targeted to the ColE1 *imm* gene. These insertions presumably were the result of

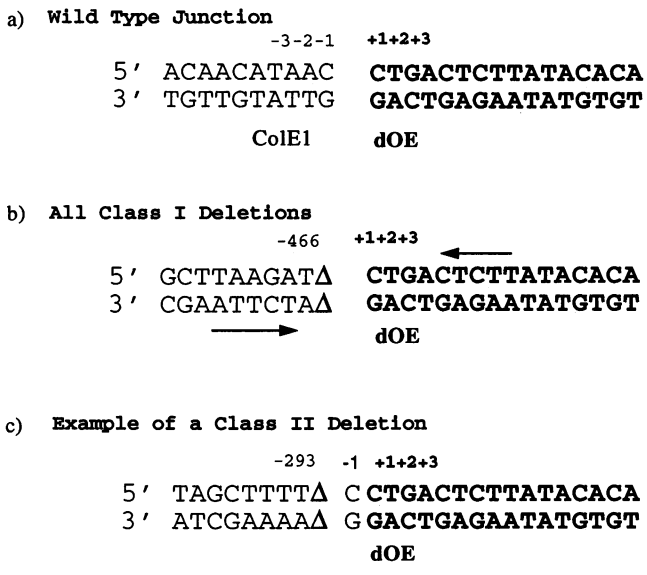

FIG. 4. (a) WT fusion junction for the dOE sequences and the neighboring ColE1 vector sequences. Each base is given a position number based on its distance from and orientation toward this junction. The dOE bases are shown in boldface type and are given progressively larger position numbers as they extend away from the junction. The vector bases are shown in lightface type and are likewise given position numbers, this time negative, to represent their distance from the junction. (b) Class I deletion junction. Using the same position numbers as those shown in panel a, one can see that class I deletions involve the deletion of 465 bases of vector DNA. The class I deletion fusion junction is always from base +1 of the dOE sequences to base −466 of the ColE1 vector sequences. We call this location in the ColE1 *imm* gene the class I deletion target site. The arrows emphasize the similarity and orientation of the target site to 5 bases of the dOE. (c) Example of a class II deletion. The start of the class II deletion fusion junction is always at base −1, although the other side of the fusion junction varies. Shown is a deletion with a fusion junction from −1 to −293.

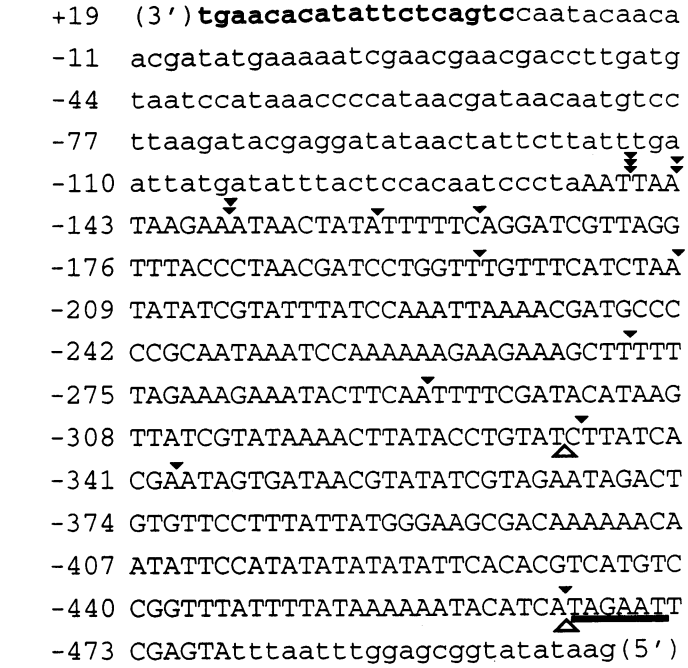

FIG. 5. Sequence of the vector next to the dOE. Each position is numbered to show its distance from the dOE. The 19 bp of the dOE are shown in boldface type. The Col E1 *imm* gene open reading frame is shown in uppercase letters. The start codon for the *imm* gene begins at position −478. The various endpoints of the deletions are shown by small solid triangles. Class II deletion endpoints that were found more than once are noted with multiple triangles. The class I deletion target site is underlined. The positions of insertional fusions are shown by large open triangles.

plasmid-to-plasmid transpositions. All 10 of the independently isolated Tn*5* insertions that we found were at one of two specific sites. Three were inserted at the class I deletion target site, and the other seven were inserted at a second site shown by large open triangles in Fig. 5. There was no apparent sequence similarity between these two insertional target sites.

DISCUSSION

Two classes of Tnp-dependent adjacent deletions were found. These two deletion types are structurally different and occur in the presence of different cOE mutations. Therefore, the cOE structure influences what happens at the dOE at a distance of 12 kbp. The fact that this influence at a distance requires the transposase protein indicates that, under normal transposition conditions, with two WT OEs, the cOE and the dOE are involved in some sort of protein-mediated interaction. Such an interaction has been seen for site-specific and transpositional recombination proteins and is considered a form of nucleoprotein synapsis (1, 13, 16).

The structural differences between class I and class II deletions suggest that they occur through different mechanisms. Below we discuss some possible mechanisms for each class.

Class I deletion formation. Class I deletions form when the cOE sequence is entirely missing (pRZ1490). We therefore assume that these deletions occur as a result of a failure of Tnp to bind to the cOE or that binding occurs in a manner such that the normal activity of Tnp at cOE is inhibited. This assumption suggests two mechanisms for their formation.

In the absence of Tnp binding to the cOE, no cOE-dOE synapsis can occur. If Tnp binds to a secondary binding site on the same DNA, such as the class I deletion target site, then perhaps secondary synapsis will occur. This synapsis may result in an attempt at excision that causes the loss of the intervening DNA (Fig. 6A). While the cutting site at the target site is 4 bp away from the normal end cutting site, this aberrant cut may be allowed by altered conformations of transposase bound at the target site. This model explains why class I deletion frequencies are enhanced in the absence of a functional binding site at the cOE. We are currently testing this model in vitro by examining the way in which Tnp binds to the different cOE mutations and to the target site. If this model is correct in its essence, it provides an interesting look at the ability of Tnp to use cryptic or mutant ends in the formation of deletions.

This model does not suggest a direct role for the various cOE mutations to determine the observed differences in the frequency of class I deletion formation. An indirect role, however, is not precluded. It is likely that some of these cOE mutations have residual binding activity of various degrees but in a conformation that is unable to yield productive synapsis. It may be that some residual binding at the mutant cOE helps to increase the deletion frequency by acting as an

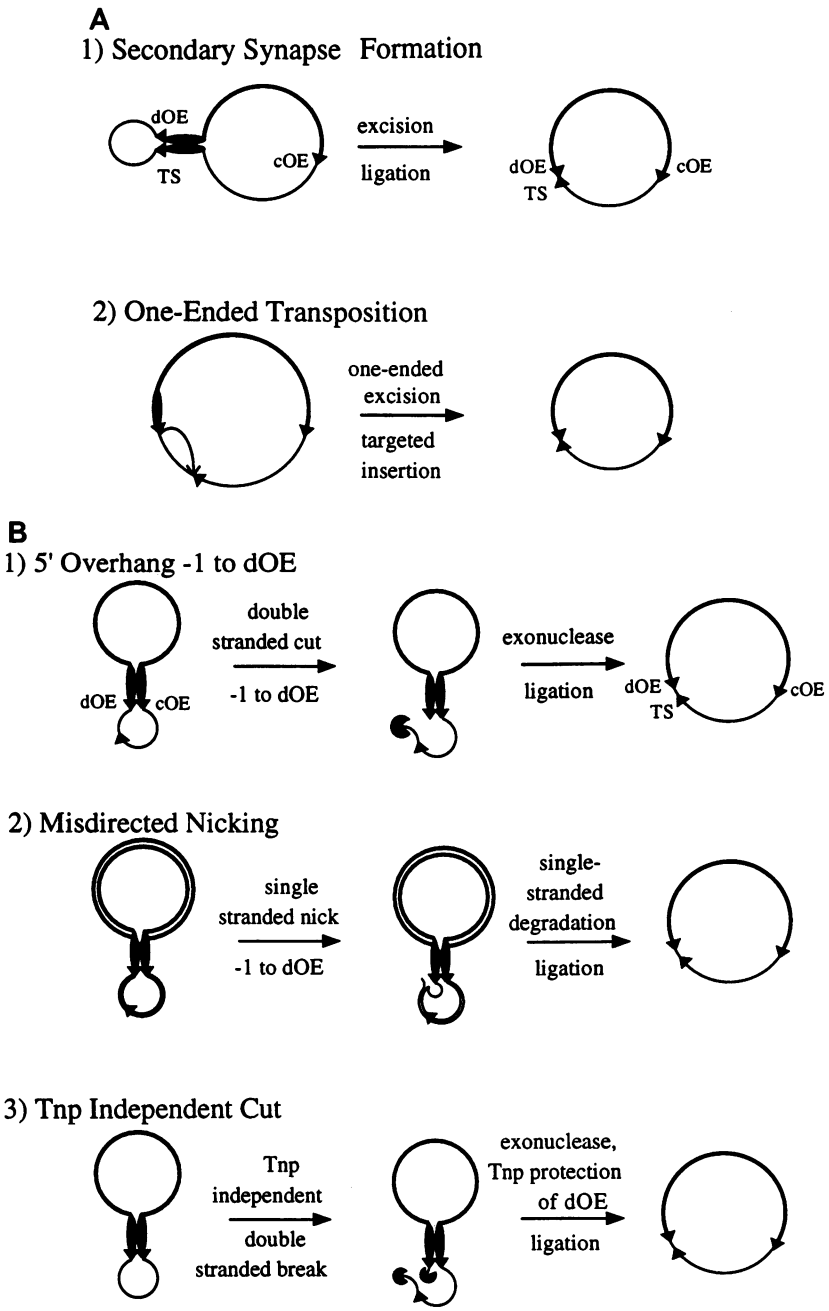

FIG. 6. Models of adjacent deletion formation. (A) Two models proposed to explain class I deletion formation. (Model 1) This model for adjacent deletion formation proposes a nucleoprotein synapsis occurring between Tnp bound at the dOE and Tnp bound at the target site (TS). The black ovals indicate bound Tnp. Excision of the intervening DNA followed by ligation results in a class I adjacent deletion. (Model 2) In the one-ended transposition model, Tnp binds initially only at the dOE. The dOE is then excised from the adjacent DNA and inserted at the TS, resulting in a class I adjacent deletion. (B) Three models proposed to explain class II deletion formation. (Model 1) The primary dOE-cOE nucleoprotein complex forms but, because of a cOE mutation, only the dOE is excised. This model proposes that Tnp normally cuts at position −1 and ligation occurs to repair the plasmid after some exonuclease activity, resulting in a class II adjacent deletion. (Model 2) This model proposes a single-strand nick position at −1 misdirected to the dOE, which results in single-strand degradation and subsequent unidirectional deletion formation. (Model 3) This model proposes Tnp-independent double-strand break of the DNA near the dOE. Exonuclease activity may occur in both directions until it reaches the dOE, which is protected by bound Tnp, and ligation occurs.

enhancer. For example, binding here may help to retain the *cis* activity of a recently made transposase molecule. This idea is consistent with the lower deletion frequency associated with the complete cOE deletion.

The fact that insertions were also found at the target site suggests an alternative model for class I deletion formation. It may be that in the presence of a nonfunctional cOE, the transposon attempts one-ended transposition. This model

involves Tnp binding only to the dOE. Once bound to the dOE, Tnp may create an excision adjacent to the bound dOE and insert the dOE into the target site. If the target site were on a different vector, such a one-ended transposition would not survive. However, if the target site were on the same vector and adjacent to the OE, then a deletion could result (Fig. 6A). Although the finding of a common insertion and deletion target site is consistent with this model, we are concerned that the other insertion target site (with the possible exception of one apparent class II deletion) is not also a deletion target site (Fig. 5). It is possible that this other, more proximal site is not a target for one-ended transposition because of steric restrictions of the DNA bending. We do not view this possibility as likely, however, as this site is 334 bp away from the dOE. The fact that an insertion target site was found is by itself interesting, however, because previously no strong target site preference was shown for Tn*5*, although some low level of specificity was reported (6).

Two variations of this one-ended transposition model are interesting and worth considering. The first uses the cOE in synapsis with the dOE, so that the dOE is activated in the one-ended transposition. Here, the cOE is unable to be cleaved but has a direct role in determining the deletion frequency on the basis of how well Tnp is able to bind the mutant end. The other variation actually uses a different, cryptic OE-like sequence internal to the dOE and target site. A synaptic complex between the dOE and the cryptic OE forms and attempts to transpose to the target site. Again, the cryptic end is not cleaved, so only the dOE is fused to the target site, resulting in one-ended transposition. A possible cryptic end is available between the dOE and the target site (3′ GAGGATAT at −86; Fig. 5). Again, an indirect influence of cOE on the deletion frequency would need to be invoked.

Class II deletion formation. Three models are proposed to explain class II deletion formation. It is possible that, although Tnp is capable of binding to both the dOE and the cOE, it only cuts at the dOE, perhaps because of the cOE mutation causing an improper binding conformation, similar to what has been reported for bacteriophage Mu (27). Some exonucleolytic degradation of the vector DNA occurs, and the remaining vector DNA is religated to the transposase-bound dOE (Fig. 6B). In this case, it is assumed that the transposase normally cuts at a position that leaves a 5′ overhang of 1 base beyond the dOE sequence. 5′ overhangs have been reported in the excision of Tn*7* (3) and several retrotransposons (10). This model does not explain why there are similar class II deletion frequencies for both WT and mutant cOEs.

A second model assumes that Tnp bound at the dOE makes a misdirected nick at the −1 position that aborts the transposition process. This nick would serve as a substrate for unidirectional deletion formation (Fig. 6b), as has been found for the DNA nicking associated with the M13 origin of replication (20, 21). cOE mutations presumably make these errors more easily detected by reducing the background of normal transposition events.

A third model proposes that a Tnp-independent double-strand break followed by exonucleolytic attack occurs somewhere near the dOE. Tnp bound at the dOE then serves merely to protect the sequence starting at position −1 until religation to the remaining vector sequence occurs (Fig. 6B). cOE mutations may enhance the probability of Tnp being bound to the dOE or, as mentioned above, may merely make these events more easily detected. This type of mechanism may explain the appearance of class III deletions in the absence of active Tnp (and therefore in the absence of dOE protection).

At this time, it is difficult to support one of these models for class II deletion formation over the others. We do note, however, that the frequencies of class II and class III deletions are similar. This result may indicate that they originate from a similar source, such as the Tnp-independent double-strand break proposed above. The fact that the 11T→A mutation gave rise to seven class I deletions and one class II deletion is interesting. It may be that class II deletions occur with all of the cOE mutations but at a frequency low enough so that they have not been observed by us.

Functional organization of the OE. Our results make it clear that there is a relationship between the cOE mutations present and the deletion class found. The similarity in the effects of class I-associated cOE mutations and the complete deletion of the cOE makes it seem likely that these base substitutions disrupt the initial binding site of the transposase or synapsis formation. This possibility suggests that positions 5 to 9 of the OE represent a region critical for protein binding. Note that the class I deletion target site has homology with positions 6 to 10 of the OE.

It appears, on the other hand, that the 2T→A and 3G→C OE mutations allow binding but may somehow interfere with a later step of the transposition process, such as excision. The 2T→C mutation, however, gives rise to class I deletions, consistent with that particular base substitution interfering with Tnp binding. The interpretation of the effects of mutations at positions 10 to 12 is less clear because the positions overlap with the DnaA binding site and because the positions are not homologous between the OE and the IE. Similar multidomain end architectures have been proposed for Tn*10* (11) and IS*903* (9). The idea that some end mutations interfere with Tnp binding while others do not is currently being tested by gel retardation techniques.

The fact that WT ends allow both classes of deletions indicates that even with intact ends, the transposition process can upon rare occasion fail and be aborted. Our results show that these nonproductive attempts at transposition may result from either a failure to form the primary synaptic complex or a failure to excise from host DNA properly. Of course, there are also likely many other critical steps necessary for a complete insertional transposition event.

ACKNOWLEDGMENTS

This work was funded by grants R01 GM19670 from the NIH, DMB9020517 from the NSF, and MV-64163 from the American Cancer Society to W.S.R.

We thank Patricia Kiley for suggestions regarding preparation of the manuscript.

REFERENCES

1. **Adzuma, K., and K. Mizuuchi.** 1989. Interaction of proteins located at a distance along DNA: mechanism of target immunity in the Mu DNA strand-transfer reaction. Cell **57**:41–47.
2. **Arthur, A., and D. Sheratt.** 1979. Dissection of the transposition process: a transposon encoded site-specific recombination system. Mol. Gen. Genet. **175**:267–274.
3. **Bainton, R., P. Gamas, and N. L. Craig.** 1991. Tn*7* transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell **65**:805–816.
4. **Berg, D. E.** 1989. Transposon Tn*5*, p. 185–210. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
5. **Berg, D. E., and M. M. Howe (ed.).** 1989. Mobile DNA. American Society for Microbiology, Washington, D.C.

6. **Berg, D. E., M. A. Schmidt, and J. B. Love.** 1983. Specificity of transposon Tn*5* insertion. Genetics **105:**813–828.
7. **Chan, P. T., H. Ohmori, J. Tomizawa, and J. Lebowitz.** 1985. Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. **260:**8925–8935.
8. **de la Cruz, N., and W. S. Reznikoff.** Submitted for publication.
9. **Derbyshire, K. M., L. Hwang, and N. D. F. Grindley.** 1987. Genetic analysis of the interaction of the insertion sequence IS*903* transposase with its terminal inverted repeats. Proc. Natl. Acad. Sci. USA **84:**8049–8053.
10. **Eichingar, D. J., and J. D. Boeke.** 1990. A specific terminal structure is required for Ty transposition. Genes Dev. **4:**324–330.
11. **Huisman, O., P. R. Errada, L. Signon, and N. Kleckner.** 1989. Mutational analysis of IS*10*'s outside end. EMBO J. **8:**2101–2109.
12. **Isberg, R. R., and M. Syvanen.** 1985. Tn*5* transposes independently of cointegrate resolution—evidence for an alternative model for transposition. J. Mol. Biol. **182:**69–78.
13. **Johnson, R. C., and M. F. Bruist.** 1989. Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombinational enhancer in the strand exchange reaction. EMBO J. **8:**1581–1590.
14. **Johnson, R. C., and W. S. Reznikoff.** 1984. Role of the IS*50* protein in the promotion and control of Tn*5* transposition. J. Mol. Biol. **176:**645–661.
15. **Krebs, M. P., and W. S. Reznikoff.** 1988. Use of a Tn*5* derivative that creates *lacZ* translational fusions to obtain a transposition mutant. Gene **63:**277–285.
16. **Lavoie, B. D., and G. Chaconas.** 1990. Immunoelectron microscopic analysis of the A,B and Hu protein content of bacteriophage Mu transposomes. J. Biol. Chem. **265:**1623–1627.
17. **Lichens-Park, A., and M. Syvanen.** 1988. Cointegrate formation by IS*50* requires multiple donor molecules. Mol. Gen. Genet. **211:**244–257.
18. **Makris, J. C., P. L. Nordmann, and W. S. Reznikoff.** 1988. Mutational analysis of insertion sequence *50* (IS*50*) and transposon *5* (Tn*5*) ends. Proc. Natl. Acad. Sci. USA **85:**2224–2228.
19. **Marinus, M. G., and E. B. Konrad.** 1973. Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. J. Bacteriol. **114:**1143–1150.
20. **Michel, B., E. D'Alançon, and S. D. Ehrlich.** 1989. Deletion hot spots in chimeric *Escherichia coli* plasmids. J. Bacteriol. **171:**1846–1853.
21. **Michel, B., and S. D. Ehrlich.** 1986. Illegitimate recombination at the replication origin of bacteriophage M13. Proc. Natl. Acad. Sci. USA **83:**3386–3390.
22. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Nordmann, P. L., J. C. Makris, and W. S. Reznikoff.** 1988. Inosine induced mutations. Mol. Gen. Genet. **214:**62–67.
24. **Phadnis, S. H., and D. E. Berg.** 1987. Identification of basepairs in the outside end of insertion sequence IS*50* that are needed for IS*50* and Tn*5* transposition. Proc. Natl. Acad. Sci. USA **84:**9118–9122.
25. **Roberts, D. E., D. Ascherman, and N. Kleckner.** 1991. IS*10* promotes adjacent deletions at low frequency. Genetics **128:**37–43.
26. **Shapiro, J. A.** 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc. Natl. Acad. Sci. USA **76:**1933–1937.
27. **Surette, M. G., T. Harkness, and G. Chaconas.** 1991. Stimulation of the MuA protein-mediated strand transfer reaction by the MuB protein, and the requirement of DNA nicking for stable type 1 transposome formation. J. Biol. Chem. **266:**3118–3124.
28. **Tomcsanyi, T., C. Berg, S. H. Phadnis, and D. Berg.** 1990. Intramolecular transposition by a synthetic IS*50* (Tn*5*) derivative. J. Bacteriol. **172:**6348–6354.
29. **Warren, G. J., M. W. Saul, and D. J. Sherratt.** 1979. ColE1 plasmid mobility: essential and conditional functions. Mol. Gen. Genet. **170:**103–107.
30. **Weinreich, M., and W. S. Reznikoff.** 1992. Fis plays a role in Tn*5* and IS*50* transposition. J. Bacteriol. **174:**4530–4537.
31. **Yin, J. C. P., M. P. Krebs, and W. S. Reznikoff.** 1988. Effect of *dam* methylation on Tn*5* transposition. J. Mol. Biol. **199:**35–45.
32. **Yin, J. C. P., and W. S. Reznikoff.** 1987. *dnaA*, an essential host gene, and Tn*5* transposition. J. Biol. Chem. **169:**4637–4645.